A theory-driven thesis: Utilising theory-driven evaluation to guide the conduct and content of a PhD thesis examining peer-led Sex and Relationships Education

Emma Dobson

School of Education, Durham University, Durham, United Kingdom

e.s.dobson@durham.ac.uk

Emma Dobson is a Postdoctoral Research Associate at Durham University's School of Education.

A theory-driven thesis: Utilising theory-driven evaluation to guide the conduct and content of a PhD thesis examining peer-led Sex and Relationships Education

The following article discusses the content and conduct of a PhD thesis exploring the utilisation of peer education to deliver Sex and Relationships Education (SRE) to adolescents in the United Kingdom. Evaluative literature currently suffers from a lack of theorisation and an absence of mechanistic investigation. In attempting to address this limitation, the PhD evolved from one study into a series of five separate studies. It is hoped that by drawing findings together from separate studies, the work as a whole will form a more complete, cohesive and comprehensive understanding of peer-led SRE. The purpose of this article is to describe how undertaking Theory-Driven Evaluation facilitated the evolution of the PhD from one proposed study to five; present an argument as to why this was the best approach to conduct the PhD study; and to outline the strengths and limitations of employing such an approach to compile the PhD thesis.

Keywords: evaluation; theory-driven evaluation; sex education; peer education

Researching peer-led SRE

One of the most popular justifications for the utilisation of young people as peer educators in Sex and Relationships Education (SRE) is the 'open communication' (Milburn, 1995) between peer educators and students. This is thought to imbue participants with 'the confidence to actively participate in discussion and to ask sensitive and intimate questions' (Fletcher, Hurst, Bolzern, & Schulkind, 2015, p.96); leading to improved knowledge and behaviour.

Discovering a theory-practice inconsistency

Despite such claims, there is limited and uneven evidence of effectiveness for peer-led SRE (Harden, Oakley, & Weston, 1999; Milburn, 1995; Stephenson et al., 2004; Stephenson et al., 2008; Tolli & Tolli, 2012) and no research investigating communicative mechanisms. Communication claims should be assessed as 'detailed analysis of social interactions between young people engaged in peer education is lacking' (Price & Knibbs, 2009, 298). This is an example of a 'theory-practice inconsistency' (Smith, 2006). There are a number of potential explanations as to why communication claims have not yet been subject to investigation (Dobson, 2016). The following article will postulate that existent evaluative literature has deterred mechanistic evaluation and proposes theory-driven evaluation as a solution to address this.

Divisive design dichotomies

Peer education research can be located in one of two academic 'camps'. Typically this is framed using language such as valuing outcomes or process (quantitative or qualitative data), effectiveness of intervention or experience of participants (positivist or constructivist). Whilst some believe that the 'paradigm wars' have abated (Arnd-Caddigan & Pozzuto, 2006), the debate is still driven by stereotypes on both sides (Gould,

2004) and appears unescapable (Maudsley, 2011). It is reminiscent of quizzes featured in *Cosmopolitan* magazine: 'Do you like numbers? If yes, go to Box 1. You are a Positivist and can only conduct outcome evaluation focussed on effectiveness'. 'If you answered no, do you like words? If yes, go to Box 2. You are a Constructivist and can only conduct process evaluation focussed on participant experience'.

This is obviously a gross simplification and caricature, but it is surprising how often peer education research positions itself in this manner. Such an approach to research is damaging. It suggests that researchers have to choose a side and collect quantitative outcome or qualitative process data based on whether they value experimental effects or participant experience. It is also a fallacy; a variety of methods and perspectives can be combined to answer research questions. Finally, it limits evidence to an 'end product'. In the case of peer education, where there is mixed evidence of effectiveness, researchers need to consider the theory at the beginning and mechanisms in the middle of an intervention, to better understand outcomes at the end.

Evaluating peer-led SRE

Evaluation of peer-led SRE has sought to establish intervention effectiveness or examine participant experience. Both goals are important, but solely focussing on these products leaves unresolved questions regarding the mechanisms at work in interventions.

Establishing effectiveness through outcome evaluation

A review of UK-based literature identified effectiveness as being assessed through either RCT or pre/post-test design. Rarely do these provide information such as where the intervention might work, for whom and under what conditions (Wrigley, 2018). Instead they aim for generalisability across large populations. This is not to dismiss RCTs. They provide the strongest evidence of causal effect (Bonell, Fletcher, Morton, Lorenc, &

Moore, 2012), but this knowledge is limited without understanding how an intervention produces its effects. 'Emphasis on measuring outcomes throws little light on the processes of learning' (Turner & Shepherd, 1999, 243). Consequently, these designs as they are currently utilised in peer education research, don't reveal which mechanisms are at work and whether these work as hypothesised.

Exploring experience through process evaluation

Most UK-based studies of peer-led SRE are process evaluations. The majority use the retrospective accounts of students, teachers and peer educators to establish acceptability or discuss technical aspects of provision such as peer educator selection and training (Forrest, Strange, & Oakley, 2002; Strange, Forrest, Oakley, & RIPPLE Study Team., 2002; Tripp, Dixon, Rees, & Kay, 2002). These are presented as evidence of effectiveness. Evaluations are typically case studies, rarely including a comparative element. Without an equivalent comparator, they cannot provide a definitive answer as to whether peer-led approaches increase enjoyment, comfort or communication in SRE compared to alternative provision (Mellanby, Rees, & Tripp, 2000), although many claim to do so.

The importance of investigating mechanisms

Current evaluative focus sheds little light on which mechanisms specifically influence intervention effectiveness or acceptability and how they do this. The practice of evaluating effects, rather than *how* effects are produced, is referred to as 'black box' evaluation. 'Much of the existing literature on peer education closely aligns with the black box approach' (Southgate & Aggleton, 2017, p.6). It rarely examines what peer educators do, or how their activities are experienced and interpreted by participants. There is nothing wrong with black box approaches to evaluation if the evaluation aim is to judge

effectiveness. Difficulties arise when products from the black box are inconsistent. Consequently, when studies of peer-led SRE are 'contradictory' (Borgia, Marinacci, Schifano, & Perucci, 2005, 514), existing research evidence doesn't suggest which factors may be responsible for this (Cornish & Campbell, 2009).

Compiling a theory-driven thesis

In the context of peer-led SRE, evaluative literature is problematic for two reasons: firstly, it does not evaluate theory. 'If a program is based on a faulty theory it will not bring about the desired change' (Astbury & Leeuw, 2010, 364). Secondly, it does not investigate mechanisms. Instead 'emphasis on measuring outcomes has led to a diminished focus on the educative processes and practices associated with peer education' (Southgate & Aggleton, 2017, 5). The original aim of the thesis was to compare different educators' communicative styles in SRE classrooms. Upon reviewing the literature however, it became clear that studies to identify and refine programme mechanisms needed to be undertaken before an experiment could be conducted.

Theory-driven evaluation

Research aims were influenced by theory-driven evaluation (TDE). The purpose of TDE is to assess the robustness of a programme's underlying assumptions (Astbury & Leeuw, 2010), focusing on intervention implementation, effectiveness, and causal mechanisms and contextual factors that facilitate or inhibit change (Chen, 1990).

TDE is recommended when:

> 'RCTs have produced inconsistent estimates of efficacy and there is no consensus on when, how and with whom to use these interventions… or when the existing research on a particular intervention is made up of mainly disparate studies and grey

> literature which do not lend themselves to statistical analysis but provide a rich source of qualitative data'
>
> (Wong, Greenhalgh, Westhorp, & Pawson, 2012, 94).

These are precisely the problems identified with the existing evidence base for peer-led SRE, making TDE a useful framework to guide the PhD study.

TDE is a general term used to describe any approach that is focussed on theory development and/or examination. As such, it can refer to an array of different evaluative approaches (Donaldson & Lipsey, 2006), the most popular being 'Theories of Change' (Weiss, 1995) and 'Realist Evaluation' (Pawson & Tilley, 1997). 'Theory-driven evaluation' was used to describe the thesis as it mixed both Theories of Change and Realist Evaluation. This balanced broad strategic learning from Theories of Change and more specific investigation through Realist Evaluation (Blamey & Mackenzie, 2007).

Thesis aim

Assumptions about open communication in the context of peer-led SRE have not been clearly articulated or evaluated. Existent literature presents some vague hypotheses about how open communication increases effectiveness; but a clear, consensual vision of this process is lacking. There is a need to develop the programme theory and mechanisms of peer-led SRE. These can be synthesised into a consistent theoretical format, producing testable hypotheses to be investigated via experimentation.

Consequently, the overarching aim of the thesis was to identify mechanisms underlying peer-led SRE to help specify and refine programme theory (Chen, 1990) to be empirically tested.

Thesis structure

Following a TDE approach, the thesis was divided into two distinct phases:

(1) Phase One: located in an inductive theoretical drive, Studies I and II examine theoretical and empirical literature. Study III explores the individual and collective practices and experiences of peer educators, practitioners and other stakeholders involved in peer-led SRE to identify programme mechanisms. Findings from these studies thereby develop a more specified and refined programme theory.

(2) Phase Two: located in a deductive theoretical drive, Studies IV and V focus on observing the presence of, and measuring changes arising from, programme mechanisms and testing these against programme effects as suggested by programme theory.

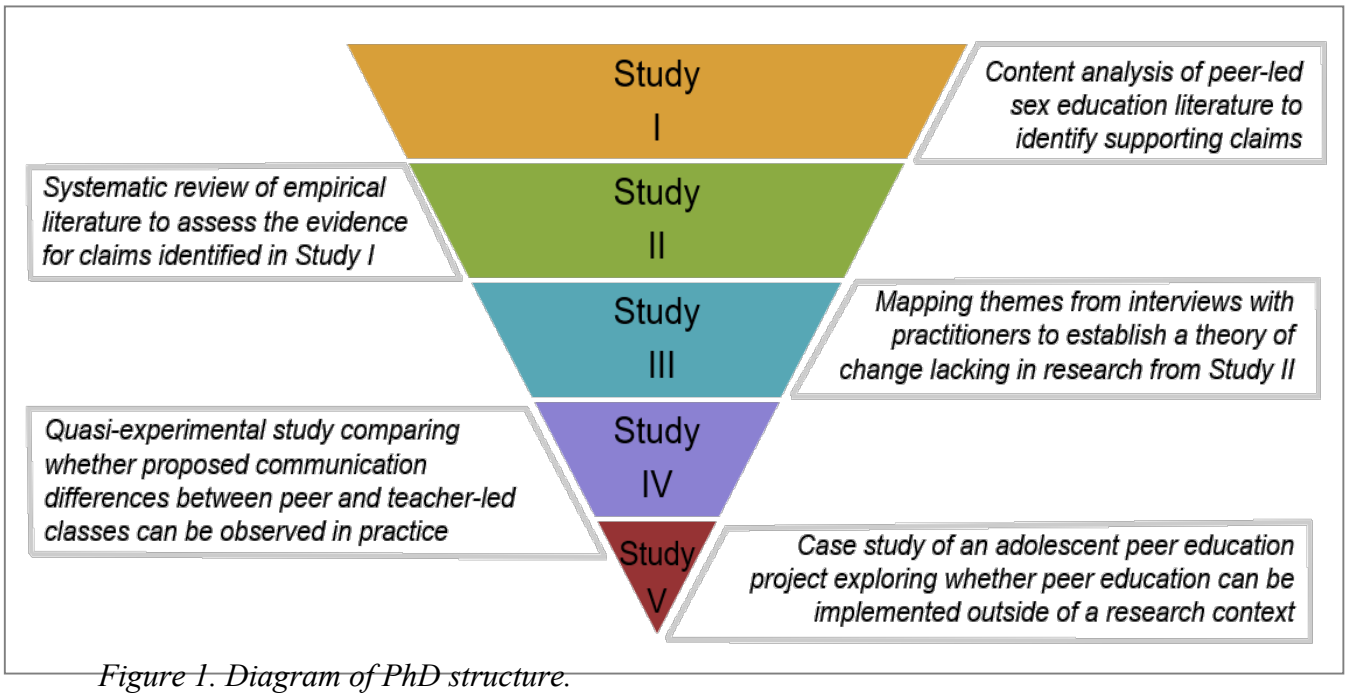

Figure 1. Diagram of PhD structure.

This structure, as depicted in Figure 1, facilitated a cumulative approach to knowledge generation and represented a retroductive narrowing of focus as the thesis progressed.

Assessing the appropriateness of a theory-driven thesis

Strengths

Adopting a TDE approach focused evaluation design by identifying research questions to test whether the proposed theory and its associated mechanisms worked as intended in practice. It also encouraged the conduct of research in different ways, as the choice of method was dictated by the research question (McEvoy & Richards, 2006). Utilising mixed methods helped to create a bridge between 'diverse perspectives on the phenomena being studied, so as to deepen, rather than simply broaden or triangulate the understanding gained' (Maxwell & Mittapalli, 2011, 147).

The approach also required a 'more intensive relationship between evaluators and key stakeholders' (Blamey & Mackenzie, 2007, 451) to uncover programme theory and the practicalities of programme delivery. Practitioner insights can be lost in research (Gough, 2004), thus it was beneficial to include their views in the research process. This may foster a sense of ownership amongst practitioners, thereby increasing engagement with research findings.

Limitations

There are no accepted quality criteria for TDE (Marchal, van Belle, van Olmen, Hoerée, & Kegels, 2012). This lack of methodological guidance is identified as a limitation when undertaking theory-driven evaluation (Rycroft-Malone, Fontenla, Bick, & Seers, 2010). This can also be a strength of the approach as its lack of prescription allows the researcher to be methodologically flexible. Despite a lack of guidance, core texts recommend evaluators start by scrutinising existing theory. A major challenge therefore is a lack of relevant theory (Chen & Rossi, 1989). This was an issue for the current thesis, but encouraged the researcher to seek out a range of literature to systematically search for

and identify programme theory, leading to a thorough understanding of the research and evidence-base.

This was a long, time-consuming process however and is an example of the most common critique of theory-driven evaluation; that it is resource and time intensive (Blamey & Mackenzie, 2007; Mackenzie & Blamey, 2005; Pedersen & Rieper, 2008).

> 'Program theory in many fields is at a low stage of development… theories that evaluations are likely to be testing are low-level approximations, riddled with inaccuracies and false paths. To discover their flaws and painstakingly revise and improve them will take multiple iterations. This is an arduous way to generate generalizable knowledge'
> (Weiss, 1998, 69).

Whilst this is a necessary caution, these efforts are more ethical than wasting funds researching or implementing interventions based on faulty or disproven theories. Considering this, it could be argued that the systematic approach of theory-driven evaluation is *more* time and cost efficient than uncoordinated efforts (Bonell et al., 2012).

Conclusion

Problematic evaluation is not specifically limited to peer-led SRE but is a common issue across disciplines such as Social Work, Education and Health. Undertaking theory-driven evaluation may be a useful approach to address this. It provides a systematic way to tease out, theorise and test mechanisms without getting lost or tied up in the various strands of the complex intervention being evaluated (Marchal et al., 2012). In this way, TDE may be of use to other PhD students when conducting their research as it helps to identify and prioritise key evaluation questions and guide the selection of data collection methods and analytical techniques. The PhD is a learning journey, and as such, should be a time in which students can experience a variety of research methods from different disciplines.

The methodological flexibility of TDE provides students with an opportunity to do this. Finally, the time required to undertake such an approach should not be a deterrent for students commencing PhD study, as they have several years of 'academic apprenticeship' to invest in their research (!).

References

Arnd-Caddigan, M., & Pozzuto, R. (2006). Truth in our time. *Qualitative Social Work*, *5*(4), 423–440.

Astbury, B., & Leeuw, F. L. (2010). Unpacking black boxes: mechanisms and theory building in evaluation. *American Journal of Evaluation*, *31*(3), 363–381.

Blamey, A., & Mackenzie, M. (2007). Theories of change and realistic evaluation: peas in a pod or apples and oranges? *Evaluation*, *13*(4), 439–455.

Bonell, C., Fletcher, A., Morton, M., Lorenc, T., & Moore, L. (2012). Realist randomised controlled trials: a new approach to evaluating complex public health interventions. *Social Science & Medicine*, *75*(12), 2299–2306.

Borgia, P., Marinacci, C., Schifano, P., & Perucci, C. A. (2005). Is peer education the best approach for HIV prevention in schools? Findings from a randomized controlled trial. *The Journal of Adolescent Health: Official Publication of the Society for Adolescent Medicine*, *36*(6), 508–516. http://doi.org/10.1016/j.jadohealth.2004.03.005

Chen, H.T., & Rossi, P. H. (1989). Issues in the theory-driven perspective. *Evaluation and Program Planning*, *12*(4), 299–306. http://doi.org/10.1016/0149-7189(89)90046-3

Chen, H. T. (1990). Issues in constructing program theory. *New Directions for Program Evaluation*, *1990*(47), 7–18.

Cornish, F., & Campbell, C. (2009). The social conditions for successful peer education: a comparison of two HIV prevention programs run by sex workers in

India and South Africa. *American Journal of Community Psychology*, *44*(1–2), 123–135.

Dobson, E. S. (2016). For successful Sex and Relationship Education, effective communication is key: but with whom? *Education & Health*, *34*(4), 91–94. Retrieved from www.sheu.org.uk/sheux/eh344finaled.pdf

Donaldson, S. I., & Lipsey, M. W. (2006). *Roles for theory in evaluation practice. Handbook of Evaluation*. Thousand Oaks, Calif: Sage.

Fletcher, J., Hurst, S., Bolzern, J., & Schulkind, J. (2015). Sexpression:UK - A new vision for comprehensive sex and relationship education. *Education & Health*, *33*(4), 96–99.

Forrest, S., Strange, V., & Oakley, A. (2002). A Comparison of Students' Evaluations of a Peer-delivered Sex Education Programme and Teacher-led Provision. *Sex Education*, *2*(3), 195–214. http://doi.org/10.1080/1468181022000025776

Gough, D. A. (2004). Systematic research synthesis to inform the development of policy and practice in education. In G. Thomas & R. Pring (Eds.), *Evidence-based Practice* (pp. 44–62). Buckingham: Open University Press.

Gould, N. (2004). Qualitative research and social work: The methodological repertoire in a practice-oriented discipline. In *Reflecting on Social Work-Discipline and Profession* (pp. 130–44). Aldershot: Ashgate.

Harden, A., Oakley, A., & Weston, R. (1999, November 1). A Review of The Effectiveness and Appropriateness of Peer-Delivered Health Promotion for Young People. Institute of Education, University of London. Retrieved from

http://eprints.ioe.ac.uk/15275/

Mackenzie, M., & Blamey, A. (2005). The Practice and the Theory. *Evaluation*, *11*(2), 151–168. http://doi.org/10.1177/1356389005055538

Marchal, B., van Belle, S., van Olmen, J., Hoerée, T., & Kegels, G. (2012). Is realist evaluation keeping its promise? A review of published empirical studies in the field of health systems research. *Evaluation*, *18*(2), 192–212. http://doi.org/10.1177/1356389012442444

Maudsley, G. (2011). Mixing it but not mixed-up: mixed methods research in medical education (a critical narrative review). *Medical Teacher*, *33*(2), e92–e104.

Maxwell, J. A., & Mittapalli, K. (2011). Realism as a stance for mixed methods research. In J. A. Maxwell (Ed.), *A Realist Approach to Qualitative Research*.

McEvoy, P., & Richards, D. (2006). A critical realist rationale for using a combination of quantitative and qualitative methods. *Journal of Research in Nursing*, *11*(1), 66–78.

Mellanby, A., Rees, J., & Tripp, J. (2000). Peer-led and adult-led school health education: a critical review of available comparative research. *Health Education Research*, *15*(5), 533–45. http://doi.org/10.1093/her/15.5.533

Milburn, K. (1995). A critical review of peer education with young people with special reference to sexual health. *Health Education Research*, *10*(4), 407–420. http://doi.org/10.1093/her/10.4.407

Pawson, R., & Tilley, N. (1997). An introduction to scientific realist evaluation. In E. Chelimsky & W. R. Shadish (Eds.), *Evaluation for the 21st Century: A handbook*

(pp. 405–418). Thousand Oaks, CA: Sage Publications. http://doi.org/http://dx.doi.org/10.4135/9781483348896.n29

Pedersen, L. H., & Rieper, O. (2008). Is Realist Evaluation a Realistic Approach for Complex Reforms? *Evaluation*, *14*(3), 271–293. http://doi.org/10.1177/1356389008090856

Price, N., & Knibbs, S. (2009). How Effective is Peer Education in Addressing Young People's Sexual and Reproductive Health Needs in Developing Countries? *Children & Society*. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1099-0860.2008.00175.x/full

Rycroft-Malone, J., Fontenla, M., Bick, D., & Seers, K. (2010). A realistic evaluation: the case of protocol-based care. *Implementation Science*, *5*(1), 38. http://doi.org/10.1186/1748-5908-5-38

Smith, M. L. (2006). Overcoming theory-practice inconsistencies: Critical realism and information systems research. *Information and Organization*, *16*, 191–211.

Southgate, E., & Aggleton, P. (2017). Peer education: From enduring problematics to pedagogical potential. *Health Education Journal*, *76*(1), 3–14.

Stephenson, J., Strange, V., Allen, E., Copas, A., Johnson, A., Bonell, C., … Oakley, A. (2008). The long-term effects of a peer-led sex education programme (RIPPLE): a cluster randomised trial in schools in England. *PLoS Medicine*, *5*(11), e224; discussion e224. http://doi.org/10.1371/journal.pmed.0050224

Stephenson, J., Strange, V., Forrest, S., Oakley, A., Copas, A., Allen, E., … The Ripple Study Team. (2004). Pupil-led sex education in England (RIPPLE study): cluster-

randomised intervention trial. *Lancet*, *364*(9431), 338–346. http://doi.org/10.1016/S0140-6736(04)16722-6

Strange, V., Forrest, S., Oakley, A., & RIPPLE Study Team. (2002). What influences peer-led sex education in the classroom? A view from the peer educators. *Health Education Research*, *17*(3), 339–349. http://doi.org/10.1093/her/17.3.339

Tolli, M. V, & Tolli, M. V. (2012, October). Effectiveness of peer education interventions for HIV prevention, adolescent pregnancy prevention and sexual health promotion for young people: a systematic review of European studies. *Health Education Research*. United Kingdom. Retrieved from http://10.0.4.69/her/cys055

Tripp, J., Dixon, F., Rees, J., & Kay, C. (2002). Peer educators experience of participation in a pedagogic sex education programme in the UK. In *14th International AIDS Conference* (pp. 113–121). Barcelona. Retrieved from http://apps.webofknowledge.com.ezphost.dur.ac.uk/full_record.do?product=WOS&search_mode=GeneralSearch&qid=3&SID=Q1JD5yJSfZ8LMLLIL4x&excludeEventConfig=ExcludeIfFromFullRecPage&page=1&doc=17

Turner, G., & Shepherd, J. (1999). A method in search of a theory: Peer education and health promotion. *Health Education Research*, *14*(2), 235–247.

Weiss, C. H. (1995). Nothing as practical as good theory: Exploring theory-based evaluation for comprehensive community initiatives for children and families. *New Approaches to Evaluating Community Initiatives: Concepts, Methods, and Contexts*, *1*, 65–92.

Weiss, C. H. (1998). *Evaluation: Methods for Studying Programs and Policies*. Upper

Saddle River, NJ: Prentice Hall.

Wong, G., Greenhalgh, T., Westhorp, G., & Pawson, R. (2012). Realist methods in medical education research: what are they and what can they contribute? *Medical Education*, *46*(1), 89–96.

Wrigley, T. (2018). The power of evidence: Reliable science or a set of blunt tools? *British Educational Research Journal*, *44*(3), 359–376.

Yin, R. (1989). *Case Study Research, Design and Methods*. California: Sage.